

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 27, 2001, 17:57:00 ; Search time 2667.65 Seconds
(without alignments)
4064.586 Million cell updates/sec

Title: US-09-784-340-3_COPY_7800_8500
Perfect score: 701
Sequence: 1 gttgtgtcagtcagagaaa.....taggattccagaaaatta 701

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1344157 seqs, 7733874588 residues

Total number of hits satisfying chosen parameters: 2688314

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl:*
1: gb_ba1:*
2: gb_ba2:*
3: gb_ba3:*
4: gb_in1:*
5: gb_in2:*
6: gb_in3:*
7: gb_om:*
8: gb_ov:*
9: gb_pat1:*
10: gb_pat2:*
11: gb_pi1:*
12: gb_pi2:*
13: gb_pi3:*
14: gb_pi4:*
15: gb_pi5:*
16: gb_pi6:*
17: gb_pi7:*
18: gb_pi8:*
19: gb_pi9:*
20: gb_pi10:*
21: gb_pi11:*
22: gb_pi12:*
23: gb_pi13:*
24: gb_pi14:*
25: gb_pi15:*
26: gb_pi16:*
27: gb_pi17:*
28: gb_pi18:*
29: gb_pi19:*
30: gb_pi20:*
31: gb_pi21:*
32: gb_pi22:*
33: gb_pi23:*
34: gb_pi24:*
35: gb_pi25:*
36: gb_pi26:*
37: gb_pi27:*
38: gb_pi28:*
39: gb_pi29:*
40: gb_pi30:*
41: gb_pi31:*
42: gb_pi32:*
43: gb_pi33:*

44: em_ov:*
45: em_pat:*
46: em_ph:*
47: em_pi:*
48: em_pi1:*
49: em_pi2:*
50: em_pi3:*
51: em_pi4:*
52: em_pi5:*
53: em_pi6:*
54: em_pi7:*
55: em_pi8:*
56: em_pi9:*
57: em_pi10:*
58: em_pi11:*
59: em_pi12:*
60: em_pi13:*
61: em_pi14:*
62: em_pi15:*
63: em_pi16:*
64: em_pi17:*
65: em_pi18:*
66: em_pi19:*
67: em_pi20:*
68: em_pi21:*
69: em_pi22:*
70: em_pi23:*
71: em_pi24:*
72: em_pi25:*
73: em_pi26:*
74: em_pi27:*
75: em_pi28:*
76: em_pi29:*
77: em_pi30:*
78: em_pi31:*
79: em_pi32:*
80: em_pi33:*
81: em_pi34:*
82: em_pi35:*
83: em_pi36:*
84: em_pi37:*
85: em_pi38:*
86: em_pi39:*
87: em_pi40:*
88: em_pi41:*
89: em_pi42:*
90: em_pi43:*
91: em_pi44:*
92: em_pi45:*
93: em_pi46:*
94: em_pi47:*
95: em_pi48:*
96: em_pi49:*
97: em_pi50:*
98: em_pi51:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
c 1	701	100.0 212904	66 AC021146	AC021146 Homo sapi
c 2	194	27.7 135350	81 AL391626	AL391626 Homo sapi
c 3	185.2	26.4 174437	60 AC007768	AC007768 Homo sapi
c 4	185.2	26.4 199702	77 AC087178	AC087178 Homo sapi
c 5	183.6	26.2 151628	64 AC016665	AC016665 Homo sapi
c 6	183.6	26.2 151874	62 AC011268	AC011268 Homo sapi
c 7	183.6	26.2 187889	65 AC019306	AC019306 Homo sapi
c 8	180.8	25.8 159930	78 AL138757	AL138757 Homo sapi

9	180.6	25.8	107484	92	HS424L16	AL034373	Human	DNA
10	180.6	25.8	168224	67	AC022976	AC022976	Homo sapi	
11	180.4	25.7	180442	71	AC068599	AC068599	Homo sapi	
12	179.4	25.6	151560	73	AC036223	AC036223	Homo sapi	
13	179.4	25.6	153038	64	AC016873	AC016873	Homo sapi	
14	178.2	25.4	178748	81	AL51493	AL51493	Homo sapi	
15	177.2	25.3	168041	75	AC073874	AC073874	Homo sapi	
16	172	24.5	186343	89	AL137248	AL137248	Human	DNA
17	172	24.5	196966	71	AC031975	AC031975	Homo sapi	
18	170.8	24.4	157367	77	AC084769	AC084769	Homo sapi	
19	170.4	24.3	180095	72	AL355507	AL355507	Homo sapi	
20	169.6	24.2	137730	92	HS365E2	AL009177	Human	DNA
21	167.6	23.9	161665	76	AC084032	AC084032	Homo sapi	
22	167.6	23.9	163120	62	AC012480	AC012480	Homo sapi	
23	166.6	23.8	169653	82	AL445430	AL445430	Homo sapi	
24	165	23.5	116841	81	AP000643	AP000643	Homo sapi	
25	165	23.5	167639	91	AP001827	AP001827	Homo sapi	
26	164.8	23.5	154268	80	AL356463	AL356463	Homo sapi	
27	164.8	23.5	157122	69	AC026235	AC026235	Homo sapi	
28	163.4	23.3	186266	91	AL512782	AL512782	Human	DNA
29	162.8	23.2	162493	75	AC073972	AC073972	Homo sapi	
30	160.8	22.9	169013	92	HS138A5	AC03818	Human	DNA
31	159.8	22.8	188351	71	AC044813	AC044813	Homo sapi	
32	158	22.5	207320	63	AC012656	AC012656	Homo sapi	
33	157.6	22.5	202793	91	CNS06C83	AL391158	Human	chr
34	157.4	22.5	184042	87	AC011235	AC011235	Homo sapi	
35	156.6	22.3	148454	86	AC007023	AC007023	Homo sapi	
36	156.6	22.3	146903	68	AC024410	AC024410	Homo sapi	
37	155	22.1	142693	62	AC012032	AC012032	Homo sapi	
38	154.8	22.1	132614	70	AC026772	AC026772	Homo sapi	
39	154.8	22.1	178340	67	AC022120	AC022120	Homo sapi	
40	154.8	22.1	198896	60	AC008658	AC008658	Homo sapi	
41	154	22.0	107587	78	AL139416	AL139416	Homo sapi	
42	154	22.0	142069	90	AL445991	AL445991	Human	DNA
43	153.6	21.9	170117	79	AL158827	AL158827	Homo sapi	
44	153.6	21.9	171575	64	AC015926	AC015926	Homo sapi	
45	153.6	21.9	182288	82	AL589843	AL589843	Homo sapi	

ALIGNMENTS

RESULT 1
LOCUS AC021146/c 212904 bp DNA HTG 07-JUL-2000
DEFINITION Homo sapiens chromosome 4 clone RP11-468M14, WORKING DRAFT
AC021146
AC021146.4 GI:8568861
VERSION AC021146.4
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 212904)
AUTHORS Waterston,R.H.
TITLE The sequence of Homo sapiens clone
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 212904)
AUTHORS Waterston,R.H.
TITLE Direct Submission
JOURNAL Submitted (14-JAN-2000) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
On Jun 16, 2000 this sequence version replaced gi:7344259.

COMMENT

Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc/index.shtml
Project Information
Center project name: H_NH0468N14
Summary Statistics

FEATURES

source

Sequencing vector: M13: 88%
Sequencing vector: plasmid: 12%
Chemistry: Dye-primer ET; 88% of reads
Chemistry: Dye-terminator Big Dye; 12% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 199729 bases at least Q40
Consensus quality: 203731 bases at least Q30
Consensus quality: 206340 bases at least Q20
Insert size: 215000; agarose-fp
Insert size: 210604; sum-of-contigs
Quality coverage: 3.75 in Q20 bases; sum-of-contigs
Quality coverage: 3.83 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 24 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1
1564: contig of 1564 bp in length
1664: gap of unknown length
1665: contig of 1421 bp in length
3085: gap of unknown length
3185: gap of unknown length
5134: contig of 1949 bp in length
5234: gap of unknown length
5235: contig of 2935 bp in length
8169: gap of unknown length
8269: gap of unknown length
8170: contig of 2072 bp in length
8270: contig of 2072 bp in length
10341: gap of unknown length
10342: contig of 3173 bp in length
13614: gap of unknown length
13714: gap of unknown length
13615: contig of 3375 bp in length
17089: gap of unknown length
17189: gap of unknown length
17090: contig of 3512 bp in length
20701: gap of unknown length
20702: contig of 4200 bp in length
25001: gap of unknown length
25002: contig of 3919 bp in length
25102: gap of unknown length
28021: contig of 4236 bp in length
29121: gap of unknown length
33356: gap of unknown length
33357: contig of 3446 bp in length
33457: gap of unknown length
36902: gap of unknown length
37002: gap of unknown length
37003: contig of 5973 bp in length
42976: gap of unknown length
43075: gap of unknown length
48339: gap of 5264 bp in length
48340: gap of unknown length
48439: gap of unknown length
48440: contig of 8090 bp in length
56529: gap of unknown length
56530: gap of unknown length
65213: contig of 8584 bp in length
65214: gap of unknown length
65314: gap of unknown length
74715: contig of 9402 bp in length
74716: gap of unknown length
74815: gap of unknown length
74816: contig of 13731 bp in length
88546: gap of unknown length
88547: gap of unknown length
88647: contig of 14721 bp in length
103367: contig of 14700 bp in length
103368: gap of unknown length
118167: contig of 14700 bp in length
118168: gap of unknown length
118267: gap of unknown length
118268: contig of 14498 bp in length
132765: gap of unknown length
132766: gap of unknown length
132865: gap of unknown length
154092: contig of 21227 bp in length
154192: gap of unknown length
154193: gap of unknown length
182144: contig of 27952 bp in length
182145: gap of unknown length
182244: gap of unknown length
182245: contig of 30660 bp in length.
Location/Qualifiers
1..212904
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="4"

BASE COUNT 69110 a 36915 c 37432 g 67121 t 2326 others
ORIGIN

Query Match 100.0%; Score 701; DB 66; Length 212904;
Best Local Similarity 100.0%; Pred. No. 7.6e-159;
Matches 701; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 gttgtgtcaagtgagaaacatggtggaacttaacacacataataataacagaac 60
|||||
DB 79805 GTCTGTGTCAGTGAGAAAACATGGAGAACTTAACCAACATAAAATAACAGAAC 79746
QY 61 agctctcttgacattctcagagaaaggttcagatcccttgtaagccactagaa 120
|||||
DB 79745 AGCTCTCTTGACCATTTCTCAGAGAAAAGAGTTCAGCATCCCTTGTAAGCCACTAGAA 79686
QY 121 gaagaaatctctcgtggaagacacatcaaccaatgaaatgagaccagaagagatg 180
|||||
DB 79685 GAAGAAATCTCTGGGAAAAGCACAATTCACCAATGAATGAGACCACAGAAAGAGTG 79626
QY 181 agggactatgtgccaaaatgtaactggtggtacccaaggtgtacctaagtgggttccaa 240
|||||
DB 79625 AGGGATCTATGTGCCAAAATGTTAACTGGATCCAGGCTGTACTAGGTGGGTTCCAA 79566
QY 241 tggggaactgtaattgttaggttaatgcaagcagagacaagtcacatggagcattctg 300
|||||
DB 79565 TGGGGAACCTTAATTTGTAGGTTAATGCAAGACGACCAAAAGTCCATGAGGCAATTCG 79506
QY 301 agactgaagatagtcactttgtgcatactgcacagaatctgactagttcaagccca 360
|||||
DB 79505 AGACTGAAGATAGTCACTTTGGCATATTCGACAGAAATCTGATCATGATTCAGGCCCA 79446
QY 361 agtagctgtatctatgttccctaataggtgtgtaccagagcagtggttaagtaaaa 420
|||||
DB 79445 AGTAGGCTGATGTAGTGTGCTATAGGGGTGTACCAAGAGGACAGTGTGTAGTAAAA 79386
QY 421 tctgtcctgacacatctgaaggaatggaagaggtgaaagatttaacgggttcagtgt 480
|||||
DB 79385 TCTGTGATGAAACATTTGAGGAATGGAAGAGAGTGTGAACATTTTAACCGTGTCACTGT 79326
QY 481 tgaataagacgtctctgtgtatgaaatctcaactatatttaaatgcatgacagac 540
|||||
DB 79325 TGACTGAAGACCTCTCTGTGTATGAAATTTCAACTATATTATTAATGATGACGAC 79266
QY 541 aaataataataataaattaccacaatagcctagttaactggttactactatt 600
|||||
DB 79265 AACATAAATTAATAAATTACACAAATGCTATGTAACTACTGGGTTTACTATT 79206
QY 601 actaagagtgaaagaaacacatcttccatttataatgaaataataataacaccta 660
|||||
DB 79205 ACTACAGAGGAGAAAAGAAACCTCATTTCCATTTTATGGAATATAATCAAAATCTCTA 79146
QY 661 taaggaagttcagagccagtagatttcagaaaaata 701
|||||
DB 79145 TAAGGAAGTTTCAGAGCCAGTAGATTTCCAGAAAAATTA 79105

RESULT 2
AL391626/c DNA HTG 20-JAN-2001
LOCUS Homo sapiens chromosome 1 clone RP11-202K23, *** SEQUENCING IN
DEFINITION PROGRESS***, 21 unordered pieces.
ACCESSION AL391626
VERSION AL391626.1 GI:9931033
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 135350)
AUTHORS Plumb,B.
TITLE Direct Submission

JOURNAL COMMENT

Submitted (20-JAN-2001) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
requests: clonerequest@sanger.ac.uk

Center: Sanger Centre
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquery@sanger.ac.uk

Project Information
Center project name: BA202K23

Summary Statistics
Assembly program: XGAP4; version 4.5
Sequencing vector: plasmid; 108752; 10% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Consensus quality: 130362 bases at least Q40
Consensus quality: 131723 bases at least Q20
Insert size: 133350; sum-of-contigs
Insert size: 144571; 2.1% error; again: 2-fp
Quality coverage: 3.68x in Q20 bases; sum-of-contigs Quality
coverage: 3.46x in Q20 bases; agarose p

NOTE: This is a 'working draft' sequence. It currently
consists of 21 contigs. The true order of the pieces
is not known and their order in this sequence record is
arbitrary. Gaps between the contigs are represented as
runs of N, but the exact sizes of the gaps are unknown.
This record will be updated with the finished sequence
as soon as it is available and the accession number will
be preserved.

1 3427: contig of 3427 bp in length
3428 3527: gap of 100 bp
3528 11200: contig of 7673 bp in length
11201 11300: gap of 100 bp
11301 14155: contig of 2855 bp in length
14156 14255: gap of 100 bp
14256 31692: contig of 17437 bp in length
31693 31792: gap of 100 bp
31793 34235: contig of 2443 bp in length
34236 34335: gap of 100 bp
34336 38905: contig of 4570 bp in length
38906 39005: gap of 100 bp
39006 43252: contig of 4247 bp in length
43253 43352: gap of 100 bp
43353 46981: contig of 3629 bp in length
46982 47081: gap of 100 bp
47082 51380: contig of 4299 bp in length
51381 51480: gap of 100 bp
51481 59247: contig of 7767 bp in length
59248 59347: gap of 100 bp
59348 66869: contig of 7522 bp in length
66870 66969: gap of 100 bp
66970 84168: contig of 17199 bp in length
84169 84268: gap of 100 bp
84269 98434: contig of 14166 bp in length
98435 98534: gap of 100 bp
98535 102586: contig of 4052 bp in length
102587 102686: gap of 100 bp
102687 105415: contig of 2729 bp in length
105416 105515: gap of 100 bp
105516 109302: contig of 3787 bp in length
109303 109402: gap of 100 bp
109403 112678: contig of 3276 bp in length
112679 112778: gap of 100 bp
112779 117961: contig of 5183 bp in length
117962 118061: gap of 100 bp
118062 128651: contig of 10590 bp in length
128652 128751: gap of 100 bp
128752 131919: contig of 3168 bp in length
131920 132019: gap of 100 bp
132020 135350: contig of 3331 bp in length.
Location/Qualifiers
1. 135350

FEATURES
SOURCE

```

/misc_feature /organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="1"
/clone="RP11-202K23"
/clone_id="RPCT-11.1"
1..3427
/note="assembly_fragment:00197
fragment_chain:1
clone_end:77
vector_side:left"
misc_feature /note="assembly_fragment:00427
fragment_chain:1"
11301..14135
/note="assembly_fragment:00358
fragment_chain:1"
14256..31692
/note="assembly_fragment:00328
fragment_chain:1"
31793..34235
/note="assembly_fragment:00316
fragment_chain:1"
34336..38905
/note="assembly_fragment:00984
fragment_chain:1"
39006..43252
/note="assembly_fragment:00304
fragment_chain:2"
43353..46981
/note="assembly_fragment:00044
fragment_chain:2"
47082..51380
/note="assembly_fragment:00086
fragment_chain:2"
51481..59247
/note="assembly_fragment:00943
fragment_chain:2"
59348..66869
/note="assembly_fragment:00672
fragment_chain:3"
66970..84168
/note="assembly_fragment:00176
fragment_chain:3"
84269..98434
/note="assembly_fragment:00450
fragment_chain:3"
98535..102586
/note="assembly_fragment:00263
fragment_chain:3"
102687..105415
/note="assembly_fragment:00058
fragment_chain:4"
105516..109302
/note="assembly_fragment:00846
fragment_chain:4"
109403..112678
/note="assembly_fragment:01261
fragment_chain:4"
112779..117961
/note="assembly_fragment:00075
fragment_chain:4"
118062..128651
/note="assembly_fragment:00056
fragment_chain:4"
128752..131919
/note="assembly_fragment:01019
fragment_chain:4"
132020..135350
/note="assembly_fragment:00544
fragment_chain:4
clone_end:SP6
vector_side:right"
BASE COUNT 42934 a 23309 c 22268 g 44830 t 2009 others

```

```

ORIGIN
Query Match 27.7%, Score 194, DB 135350, Length 135350;
Best Local Similarity 64.0%, Pred. No. 1,36-1,36;
Matches 379; Conservative 0; Mismatches 5; Indels 18; Gaps 5;

QY 1 ggtgtgtcgaagtgtgaaacatctggagaaacttaccacacacataaatacagaac 60
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 86551 GGTGTGTCAGATGATGATACATAGAGCAACTCATTAACACATTAATGACAGAAC 86492

QY 61 agtctcttgaccattct-agagaaagggttcagcatccttgtaagccactatg 119
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 86491 AGTTATGACATACAGGCTGTGAGAGAAACAGAGCAACCTTGCGGGTTAATGCGA 86432

QY 120 agaagaaattctctgggaaagacatcacaacaaatgaagagacaaagaagagc 179
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 86431 AGTGGAGGCCATCTTGGACATATGATCAATCAGTAGAGGAGCAAG-AGAGAGA 86374

QY 180 gagggatctatgtgccaaatgttaactgggagtcagggtgttacctagtggttcca 239
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 86373 GGGGTCCTGTGACCCAAAGCCCTTACTGGGTCAGGCGATATGACAGGTTTCTT 86314

QY 240 atgggaactgtaattgttaggttaatgaagcaagcagaagtcacatgagagcattct 299
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 86313 GCAGGAGTTTAAATVGTAGGTTAGACAAAGACATVGTCTCCATVAGGAGCAGGCT 86254

QY 300 gagactgaaagtagtcaacttggtcatatctgacagaatctgcagtgatt----- 352
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 86253 GTGACTGAGAGGTGTGCTCATGATGATATTTGCTCCATVAGGAGGAGTATATCAATG 86194

QY 353 --caagcccaagtagctatctatgtgtccata----ygtgtgtaccagaagagca 405
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 86193 GGACAAGTCCAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 86134

QY 406 gtgtgtaagtaaaatcctgactgaacacattgaggaatgaaggaaggtgaaatttt 465
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 86133 GATGCACAGCAAAATGTGATGATGATGATGATGATGATGATGATGATGATGATG 86074

QY 466 aaaggtgtcagtggttaagacgtgtctgtgtgtatgaaatcaactatattt 524
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 86073 AAACGTGTCAAAAGGTGACCATACCTTCTTGTAAGAGACAGTCCAGCATATTTA 86014

QY 525 aatgcatagccagacacataaaattatagaattaccacaatagctatg 576
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 86013 AATGATCTCTGAGCAATATAATTTAAGATTTGCTGCTGCTGATG 85962

RESULT 3
AC007768 174437 bp DNA HTG 20-APR-2000
LOCUS Homo sapiens chromosome 18 clone RP11-540M4 map 18, WORKING DRAFT
DEFINITION SEQUENCE, 9 unordered pieces.
ACCESSION AC007768.2 GI:7622399
VERSION AC007768.2
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 174437)
AUTHORS Birren,B., Linton,L., Nusbaum,C. and Lander,E.
TITLE Homo sapiens chromosome 18, clone RP11-540M4
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 174437)
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,
Baker,J., Balow,J., Baines,N., Beckwith,R., Benn,J., Brown,A.,
Castle,A., Cerny,J., Colangelo,M., Collins,S., Collymore,A.,
Cooke,P., Dearlano,K., Depayre,E., Devon,K., Dewar,K.,
Donlan,L., Doyle,M., Ferreira,P., Fitzhugh,W., Forrest,C.,
Funke,R., Gage,D., Galagan,J., Gardy,S., Gilbert,D., Grant,G.,
Haggs,B., Heath,A., Horton,L., Howland,J.C., Jones,C., Kahn,L.,
Karalas,A., Lenocky,J., Lieu,C., Locke,K., Macdonald,P.,

```



```

SOURCE      human.
ORGANISM    Homo sapiens
REFERENCE    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS     Mammalia; Euteria; Primates; Catarrhini; Homnidae; Homo.
TITLE       1 (bases 1 to 199702)
JOURNAL     Waterston, R.H.
REFERENCE   The sequence of Homo sapiens clone
AUTHORS     Unpublished
TITLE       2 (bases 1 to 199702)
JOURNAL     Waterston, R.H.
REFERENCE   Direct Submission
AUTHORS     Submitted (11-DEC-2000) Genome Sequencing Center, Washington
TITLE       University School of Medicine, 4444 Forest Park Parkway, St. Louis,
JOURNAL     MO 63108, USA
COMMENT     On Jan 13, 2001 this sequence version replaced gi:11612642.

----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc/index.shtml
----- Project Information -----
Center project name: H_NH0793A13
----- Summary Statistics -----
Sequencing vector: M13; 100%
Sequencing vector: Plasmid; 0%
Chemistry: Dye-terminator Big Dye; 0% of reads
Chemistry: Dye-terminator Big Dye; 0% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 174880 bases at least Q40
Consensus quality: 182504 bases at least Q30
Consensus quality: 186170 bases at least Q20
Insert size: 199000; agarose-fp
Insert size: 194009; sum-of-contigs
Quality coverage: 2.87 in Q20 bases; agarose-fp
Quality coverage: 3.01 in Q20 bases; sum-of-contigs
----- NOTE: This is a 'working draft' sequence. It currently
* consists of 41 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1      1267: contig of 1267 bp in length
*      1368      1367: gap of unknown length
*      1368      2400: contig of 1033 bp in length
*      2401      2500: gap of unknown length
*      2501      4656: contig of 2156 bp in length
*      4657      4756: gap of unknown length
*      4757      6882: contig of 2126 bp in length
*      6883      6982: gap of unknown length
*      6983      8935: contig of 1953 bp in length
*      8936      9035: gap of unknown length
*      9036      11452: contig of 2417 bp in length
*      11453      11552: gap of unknown length
*      11553      13771: contig of 2219 bp in length
*      13772      13871: gap of unknown length
*      13872      15587: contig of 1716 bp in length
*      15588      15687: gap of unknown length
*      15688      19023: contig of 3336 bp in length
*      19024      19123: gap of unknown length
*      19124      22249: contig of 3126 bp in length
*      22250      22349: gap of unknown length
*      22350      24092: contig of 1743 bp in length
*      24093      24192: gap of unknown length
*      24193      26788: contig of 2596 bp in length
*      26789      26888: gap of unknown length
*      26889      30112: contig of 3224 bp in length
*      30113      30212: gap of unknown length
*      30213      32726: contig of 2514 bp in length
*      32727      32826: gap of unknown length
*      32827      36374: contig of 3548 bp in length

```

```

*      36375      36474: gap of unknown length
*      36475      39233: contig of 2759 bp in length
*      39233      39333: gap of unknown length
*      39334      42723: contig of 3390 bp in length
*      42724      42823: gap of unknown length
*      42824      45983: contig of 3160 bp in length
*      45984      46083: gap of unknown length
*      46084      48934: contig of 2851 bp in length
*      48935      49034: gap of unknown length
*      49035      54923: contig of 5869 bp in length
*      49036      55023: gap of unknown length
*      55024      59263: contig of 4240 bp in length
*      59264      59363: gap of unknown length
*      59364      63218: contig of 3855 bp in length
*      63219      63318: gap of unknown length
*      63319      67133: contig of 3815 bp in length
*      67134      72233: gap of unknown length
*      72234      72333: contig of 5000 bp in length
*      72334      79233: gap of unknown length
*      79234      79333: contig of 6900 bp in length
*      79334      79333: gap of unknown length
*      86110      86210: contig of 6777 bp in length
*      86211      92685: contig of 6475 bp in length
*      92686      92785: gap of unknown length
*      92786      98826: contig of 6041 bp in length
*      98827      98926: gap of unknown length
*      98927      105081: contig of 6155 bp in length
*      105082      105181: gap of unknown length
*      105181      110628: contig of 5447 bp in length
*      110629      110728: gap of unknown length
*      110729      118730: contig of 8002 bp in length
*      118731      118830: gap of unknown length
*      118830      126388: contig of 7558 bp in length
*      126389      126489: gap of unknown length
*      126489      134257: contig of 7769 bp in length
*      134258      134357: gap of unknown length
*      134358      141443: contig of 7066 bp in length
*      141444      141543: gap of unknown length
*      141544      150881: contig of 9338 bp in length
*      150881      150981: gap of unknown length
*      150982      160811: contig of 9830 bp in length
*      160812      160911: gap of unknown length
*      160912      170464: contig of 9553 bp in length
*      170465      170564: gap of unknown length
*      170565      184947: contig of 14383 bp in length
*      184948      185047: gap of unknown length
*      185048      197809: contig of 12762 bp in length
*      197810      197909: gap of unknown length
*      199041      199041: contig of 1132 bp in length
*      199042      199141: gap of unknown length
*      199142      199702: contig of 561 bp in length.

FEATURES
source
1. .199702
   /organism="Homo sapiens"
   /db_xref="taxon:9606"
   /chromosome="RP11-793A13"
   /clone="RP11-793A13"
1. .1267
   /note="assembly_name:Contig12"
1368. .2400
   /note="assembly_name:Contig13"
2501. .4656
   /note="assembly_name:Contig14"
4757. .6882
   /note="assembly_name:Contig15"
6983. .8935
   /note="assembly_name:Contig16"
9036. .11452
   /note="assembly_name:Contig17"
11553. .13771
   /note="assembly_name:Contig18"
13872. .15587

```

Query Match	26.4%	Score 185.2	DB 77	Length 199702
Best Local Similarity	64.0%	Prod. No. 1.7e-34		
Matches 362	Conservative	0	Mismatches 193	Indels 11
			Gaps	
QY	10	aagtgagaaacatgaggagaacttaaccacaacataataacagaacagcttcctt	69	
Db	190993	AAAAAACACACAGAGGAGTCACTCAACCAAAACACGTGAATTAAGTATTACTACT	1910523	
QY	70	tgaccattctbagaagaaagagttcagcatcccttgytaagccactgagagaagaaat	129	
Db	191053	TGACAGATCCACAGAGAGAGAGGTCAACATGCTTGGAAACCAACAAATGAAAGGGGAGT	1911122	
QY	130	tctcttgagaaagacacattcaaccaatgaaatggaagccaaagaaagagtgaaggatcta	189	
Db	191113	CATPAAGC-CATTGCCAACACAGCGGTGAAGAAATAAAGGGAGACAAAAATGAGGGACCTTG	1911707	
QY	190	tgtagcacaatatgttaacttggatcttcaggggtgttacctaggttggtgtttccaatggggaact	249	
Db	191171	TAGGCTGAAGCTTTATTATGGGGTCCA-GGTATCATCTAAGGTGGGTTCCATTGGGAATT	1912299	
QY	250	gtaattgtaggttttaatgcagaacagccacaagttcatatgagagcatctgagacgtgaaa	309	
Db	191230	GTAACGTGTGGATTTAAACGCAAGCTGGATCAAGTTCATTCGATCAACACGTGTGACTTAAA	1912899	

RESULT	5
LOCUS	AC016665/c
DEFINITION	AC016665 151628 bp DNA Homo sapiens clone RP11-5G23, WORKING DRAFT SEQUENCE, 3 unordered pieces.
ACCESSION	AC016665
VERSION	AC016665.2 GI:7230027
KEYWORDS	HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE	human.
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS	1 (bases 1 to 151628) Birren,B., Linton,L., Nusbaum,C. and Landier,E.
TITLE	Homo sapiens clone RP11-5G23
JOURNAL	Unpublished
REFERENCE	2 (bases 1 to 151628)
AUTHORS	Birren,B., Linton,L., Nusbaum,C., Landier,E., Allen,N., Anderson,M.

TITLE	Direct Submission
JOURNAL	Submitted (04-DEC-1999) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT	On Mar 12, 2000 this sequence version replaced gi:5524221.

```

http://ftp.genome.washington.edu/RM/R_speethasker.html
-----
Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: MIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence.submissions@genome.wi.mit.edu
-----
Project Information
Center project name: L2854
Center Clone name: 5_G-23
-----
Summary Statistics
Sequencing vector: M13; M7781; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 139851 bases at least Q40
Consensus quality: 147726 bases at least Q30

```

5.

* is not known and their order in this sequence record is

* arbitrary. Gaps between the contigs are represented as

Db 93191 CATAAGG--CATTC CCAACCAGCGGTAA GGAATAAGAGGAGACAAATGAGGACCTG 93134

* runs of N, but the exact sizes of the gaps are unknown.


```
repeat_region      2716..2881
                    /note="AluSc repeat: matches 1..109 of consensus"
misc_feature        /note="complement(3169..3588)
                    /note="match: GSS: Em:AQ210672"
repeat_region      3593..3677
                    /note="MIR repeat: matches 2..92 of consensus"
repeat_region      5356..5502
                    /note="MIR5B repeat: matches 3..178 of consensus"
repeat_region      5729..5917
                    /note="L2 repeat: matches 2533..2728 of consensus"
repeat_region      7136..7236
                    /note="MIR repeat: matches 46..172 of consensus"
repeat_region      8195..8495
                    /note="L1PA15 repeat: matches 5057..5357 of consensus"
repeat_region      8494..9287
                    /note="L1PA15 repeat: matches 5362..6155 of consensus"
repeat_region      9299..9439
                    /note="FLAM_A repeat: matches 17..116 of consensus"
repeat_region      9583..9700
                    /note="L2 repeat: matches 2626..2735 of consensus"
repeat_region      9874..10101
                    /note="L2 repeat: matches 2478..2706 of consensus"
repeat_region      10334..10369
                    /note="L2 repeat: matches 2478..2706 of consensus"
repeat_region      12720..13007
                    /note="L2 repeat: matches 2478..2706 of consensus"
repeat_region      13803..13923
                    /note="AluY repeat: matches 1..311 of consensus"
repeat_region      13924..14203
                    /note="MSTC repeat: matches 291..405 of consensus"
repeat_region      14204..14467
                    /note="AluSg repeat: matches 1..286 of consensus"
repeat_region      14475..16120
                    /note="MSTC repeat: matches 3..291 of consensus"
repeat_region      16128..16486
                    /note="MSTC-internal repeat: matches 5..1574 of consensus"
repeat_region      16558..16779
                    /note="MSTC repeat: matches 3..403 of consensus"
repeat_region      17174..17705
                    /note="MIR repeat: matches 34..262 of consensus"
repeat_region      21703..22157
                    /note="MIR67D repeat: matches 1..514 of consensus"
repeat_region      22162..22402
                    /note="L1R33 repeat: matches 57..521 of consensus"
repeat_region      22403..22599
                    /note="L1MA8 repeat: matches 3984..4215 of consensus"
repeat_region      23832..24478
                    /note="AluSg repeat: matches 103..298 of consensus"
repeat_region      25127..25188
                    /note="L1MA8 repeat: matches 5622..6291 of consensus"
repeat_region      25208..25325
                    /note="MIR repeat: matches 260..320 of consensus"
repeat_region      26113..26260
                    /note="MIR repeat: matches 125..262 of consensus"
repeat_region      26385..27037
                    /note="MIR repeat: matches 33..195 of consensus"
repeat_region      27902..28118
                    /note="L1ME1 repeat: matches 5513..6168 of consensus"
repeat_region      28119..28458
                    /note="match: GSS: Em:B94499"
repeat_region      28982..29088
                    /note="AluJo repeat: matches 1..309 of consensus"
repeat_region      29420..29627
                    /note="L2 repeat: matches 2640..2750 of consensus"
repeat_region      29628..29937
                    /note="MIR11 repeat: matches 161..447 of consensus"
repeat_region      29938..30035
                    /note="AluSx repeat: matches 3..310 of consensus"
repeat_region      31367..31823
                    /note="MIR11 repeat: matches 69..161 of consensus"
repeat_region      31824..32117
                    /note="L1ME2 repeat: matches 5702..6155 of consensus"
repeat_region      32118..32211
                    /note="AluSx repeat: matches 1..294 of consensus"

/note="L1ME2 repeat: matches 5609..5702 of consensus"
repeat_region      32215..32280
                    /note="L1M4 repeat: matches 5236..5298 of consensus"
repeat_region      32333..32444
                    /note="L1M4 repeat: matches 5127..5244 of consensus"
repeat_region      33471..33644
                    /note="MIR5B repeat: matches 2..178 of consensus"
repeat_region      34689..34787
                    /note="MIR repeat: matches 158..262 of consensus"
repeat_region      39859..40043
                    /note="MIR repeat: matches 11..195 of consensus"
repeat_region      40543..41568
                    /note="L2 repeat: matches 1574..2749 of consensus"
repeat_region      43171..43471
                    /note="AluSx repeat: matches 2..303 of consensus"
repeat_region      43738..44056
                    /note="AluYb8 repeat: matches 1..306 of consensus"
repeat_region      44810..48552
                    /note="L2 repeat: matches 1826..2191 of consensus"
repeat_region      48553..49095
                    /note="MIR repeat: matches 1..364 of consensus"
repeat_region      49564..49856
                    /note="L2 repeat: matches 1162..1826 of consensus"
repeat_region      50011..50901
                    /note="AluSg repeat: matches 1..301 of consensus"
repeat_region      50902..50979
                    /note="MIR54A repeat: matches 3..902 of consensus"
repeat_region      51607..51700
                    /note="L2 repeat: matches 1..107 of consensus"
repeat_region      51917..52023
                    /note="L1P1 repeat: matches 6060..6151 of consensus"
repeat_region      52226..53427
                    /note="L1MB6 repeat: matches 4969..6167 of consensus"
repeat_region      54034..54687
                    /note="L2 repeat: matches 2003..2750 of consensus"
repeat_region      55851..55884
                    /note="L17 copies 2 mer tg 82% conserved"
repeat_region      57842..57986
                    /note="MIR repeat: matches 7..234 of consensus"
repeat_region      59174..59318
                    /note="L2 repeat: matches 1823..1966 of consensus"
repeat_region      61376..61471
                    /note="L2 repeat: matches 1254..2316 of consensus"
repeat_region      61556..61645
                    /note="L2 repeat: matches 2234..2316 of consensus"
repeat_region      61646..61754
                    /note="L2 repeat: matches 2234..2316 of consensus"
repeat_region      61755..62058
                    /note="MIR63B repeat: matches 328..436 of consensus"
repeat_region      62059..62387
                    /note="AluJo repeat: matches 3..300 of consensus"
repeat_region      62388..62479
                    /note="MIR63B repeat: matches 1..328 of consensus"
repeat_region      62480..62951
                    /note="L2 repeat: matches 2316..2409 of consensus"
repeat_region      62952..62814
                    /note="match: GSS: Em:AQ039222"
repeat_region      62831..62866
                    /note="MIR53 repeat: matches 1..168 of consensus"
repeat_region      63054..63374
                    /note="L2 repeat: matches 1..294 of consensus"

Query Match      25.8%; Score 180.6; DB 92; Length 107484;
Best Local Similarity 64.7%; Pred. No. 2; Le-33;
```



```

/db_xref="taxon:9606"
/chromosome="7"
/map="7"
/clone="RP11-516P3"
/clone_lib="RPC1-11 Human Male BAC"
1..1082
/misc_feature
/feature="assembly-fragment"
1183..3006
/misc_feature
/feature="assembly-fragment"
3107..4880
/misc_feature
/feature="assembly-fragment"
4981..7611
/misc_feature
/feature="assembly-fragment"
7712..11440
/misc_feature
/feature="assembly-fragment"
11541..14838
/misc_feature
/feature="assembly-fragment"
14939..19431
/misc_feature
/feature="assembly-fragment"
19532..23497
/misc_feature
/feature="assembly-fragment"
23598..332704
/misc_feature
/feature="assembly-fragment"
332805..42394
/misc_feature
/feature="assembly-fragment"
clone_end:17
vector_side:right"
42495..53530
/misc_feature
/feature="assembly-fragment"
53631..65493
/misc_feature
/feature="assembly-fragment"
65594..82936
/misc_feature
/feature="assembly-fragment"
83037..104574
/misc_feature
/feature="assembly-fragment"
clone_end:5P6
vector_side:right"
104675..127872
/misc_feature
/feature="assembly-fragment"
127973..168224
/misc_feature
/feature="assembly-fragment"
52574 a 30015 c 30383 g 53746 t 1506 others
ORIGIN

```

Query Match	25.8%	Score 180.6	DB 67	Length 16822
Best Local Similarity	64.7%	Pred. No. 2.2e-33		
Matches 381	Conservative 0	Mismatches 194	Indels 14	Gaps 7
QY 1	gtgtgtgtcgaattgagaacaatgaggagaacttaaccacaacataataacagaac	60		
DB 60167	GTGTCAATCACTGAGACACAGAGGAGGCACACATATAACACATGAAATACAGAAC	60226		
QY 61	agct-tctttgacattctagagaaaaagtltaagcatcccttgtaagccataga	119		
DB 60227	AGTGTCATCACTATGATCTCTGCGGAGAAAGGGCGGCTACTCTCCAAAGGCCAAATGGA	60286		
QY 120	agaagaaaattctcgggagaagaacattcaaccatgtatgtagagaca--agaagaga	177		
DB 60287	ACCTGGAGACTCTCTTGAATAACAGCTGCACACTGCTGTGGAGAGCAGAGAGAGAA	60346		
QY 178	gtgaggaatcctatgtgccaaaatgttaactctggatctcaagtggttaaccaggtaggttc	237		
DB 60347	GTGAGGAGACTCTATTAGGCCAGACCCCTTAACTGGTGTCCAGGCTTAAACACAGAGGTTT	60406		
QY 238	caatgggaaactgttaattgtagt----ttaatgaagcaggcacaaatccatggag	293		
DB 60407	CCTCATGAGAGTTCTTAATTTGGTCAAGTGAATGAGACAAAGGGGCACTACTCAGTGGACT	60466		
QY 294	cattcttgagactgaaagaatagtcactcttgagatctgtgacaa-gaatctgacagat	352		
DB 60467	CATGTGTG-TACGAGAGGGTGTCCCTGTGGCATATCTTGGGAGAGTGTGGGGTCAAGG	60525		

OY	353	caagcccaagtagagctgfatctatcgaatgctgacataagggc	gttaccagaagaccag	408
Db	60526	GCAATTCACATATCTCTTGTAACAGCTGTGCCATGGAAAT	TTTCACACAGGACCACTT	60585
OY	409	tgtaatglaaaaatccctgtagctgaacacatctgggaatgtg	ttgaggtggaattttaa	468
Db	60586	GTATGAGGTAGATCTTGATTGATTCACCTTAAGGACAGG	ATAGAGCGAGAGCAACTGCAG	60645
OY	469	cgggtgcaggtgttgactaagaacctgaccttcgtatg- ^{gac}	tttcaacttaattttaa	527
Db	60646	CTGTGCCAATGGTTGCTTAACCTTCGCTTCTGTTAGAGCA	ATTCAATTTATATTTAAA	60705
OY	538	tgccttaggcagacacataaaatlaaagaattaccacaaatagctatg		576
Db	60706	TGGATGCCAAGGTACATTAATTTATAGAAATTCCTGCAGCTGTGTG		60754

RESULT	11
AC068599	
LOCUS	
DEFINITION	AC068599 180442 bp DNA HTG 07-DEC-2000 Homo sapiens chromosome 2 clone RP11-548J11, WORKING DRAFT
ACCESSION	AC068599
VERSION	AC068599.3 GI:11597175
KEYWORDS	HTG; HTGS_PHASEL; HTGS_DRAFT.
SOURCE	human.
ORGANISM	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE	1 (bases 1 to 180442)
AUTHORS	Waterston,R.H.
TITLE	The sequence of Homo sapiens clone
JOURNAL	Unpublished
REFERENCE	2 (bases 1 to 180442)
AUTHORS	Waterston,R.H.
TITLE	Direct Submission
JOURNAL	Submitted (04-MAY-2000) Genome Sequence Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis MO 63108, USA
COMMENT	On Dec 7, 2000 this sequence version r placed gi:8570302.

```

----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc/index.shtml
----- Project Information -----
Center project name: H.NH0549j11
----- Summary Statistics -----
Sequencing vector: M13, 95%
Sequencing vector: plasmid, 5%
Chemistry: dye-primer ET, 95% of reads
Chemistry: dye-terminator Big Dye, 5% of reads
Assembly program: Phrap, version 0.990-19
Consensus quality: 174697 bases at least Q40
Consensus quality: 176252 bases at least Q30
Consensus quality: 177158 bases at least Q20
Insert size: 177000; agarose-fp
Insert size: 179542; sum-of-contigs
Quality coverage: 5.02 in Q20 bases; agarose-fp
Quality coverage: 5.41 in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 10 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
1      2460: contig of 2460 bp in length
*
2461      2560: gap of unknown length
*
2561      8693: contig of 6133 bp in length
*
8694      8793: gap of unknown length

```



```

----- Summary Statistics
Center Clone name: 362_A_12
Sequencing vector: M13; M7815; 48% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 151269 bases at least Q40
Consensus quality: 151406 bases at least Q30
Consensus quality: 151491 bases at least Q20
Insert size: 15000; agarose-fp
Insert size: 151560; sum-of-ctrls
Quality coverage: 10.6 in Q20 bases; agarose-fp
Quality coverage: 10.7 in Q20.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 1 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
* 1 151560: contig of 151560 bp in length.
assembly-fragment.
Location/Qualifiers
source 1. 151560
/organism="Homo sapiens"
/db_xref="taxon:9606"
/map="2"
/clone="RP11-362A12"
/clone_lib="RPC1-11 Human Male BAC"
BASE COUNT 44866 a 27588 c 28440 g 50665 t 1 others
ORIGIN

```

```

Query Match 25.6%; Score 179.4; DB 71; Length 151560;
Best Local Similarity 63.4%; Pred. No. 4.2e-33;
Matches 381; Conservative 0; Mismatches 196; Indels 24; Gaps 6;

QY 1 ggtgtgtaagtgagagaacacatgaggaacttaacacataaataacagaac 60
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 13383 GTGTCACTCAGGAGACACAGAGACACACACAAACATGAAATACAGAGC 13442

QY 61 agctctt-tgaccattttagagaaagttcagcctccttgtaagccactgga 119
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 13443 AGTTTATTACTTACAGATTCATGAGAGAGGACACCTTCACAGAGCCCATGGGA 13502

QY 120 agagaagaattctctggaagaagacacatcaacgaatgagacacagaaga 175
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 13503 AG-GGAGAGCTTTCACACATACACACTTACACAGTGTGATGAGGCAAGAGAGGTG 13561

QY 176 gattgaggaattcattgtccaaaatgttaactggatccaggggtgttaacctagtggtt 235
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 13562 GTGCGGCTACTCTGTGTAACCTGGGTCTTTATTAGGGTCCAGAGATTGCTTAGGTGAGTT 13621

QY 236 tccaatgaggaactgtaattggttaattgaagcaagcacaagaagtcacatgagga 295
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 13622 TCTTTGTGGGAGTTCTTAATTCATGAGTTTAAACAAAGCTCTCTGAGTCTGTGGAGTCA 13681

QY 296 ttctgagactgaagatagctactttgcatatctgacagaaatct-----ga 343
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 13682 CACTGTGCGTGGAGAAATGTTACTGTGATGATCTGGGCAATGTCATGATGAGTGGGGA 13741

QY 344 tcaatgagactgaagcagaatgagctgtatctgctcctataggggtgttaccagaagg 403
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 13742 TAAAGTGGAGAAATTCAGAGAGCTTGTATCTAGCTGCACATAGGAGGTGTCACATAGG 13801

QY 404 cagtgtgtaagtaaaatcctgactgaa-----cacatgaggaagaatgaggaagg 458
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 13802 AGGTGCTGTGTGAGACAAATGTCTGTACACACACACTGAAATCTGGAGAGGGGTGAA 13861

QY 459 agatttaacaggtgtcagtggttgaactgaagactgtcttctgtaagtaaaaa-ttcaactt 517

```

```

Db 13862 GAACGTGAAACCCGTCAGAGACAACTAATCAGCTCTTGTGAGATAGAAAAGTAAACCT 13921
QY 518 atatttaatgcatgaccagacacataaattataagaattacacaaatgctg 577
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 13922 ATATTGAAATGCGTGTCAAGACAACTAATTAATGAAAATTCACATCAAAAGCTATTG 13981
QY 578 t 578
Db 13982 T 13982

RESULT 13
AC016873/c
LOCUS
DEFINITION
HOMO SAPIENS CLONE RP11-19B22, WORKING DRAFT SEQUENCE, 10 UNORDERED
PIECES.
ACCESSION
AC016873 GI:7381819
VERSION
AC016873.3
KEYWORDS
HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 153038)
Birren,B., Linton,L., Nusbaum,C. and Lander,E.
Homo sapiens, clone RP11-19B22
Unpublished
2 (bases 1 to 153038)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,
Baldwin,J., Barna,N., Beckerly,R., Bonuslavsky,L., Bouhgalter,B.,
Brown,A., Castle,A., Colangelo,M., Collins,S., Collymore,A.,
Cooke,P., Dearlano,K., Dewar,K., Domingo,M., Donelan,L., Doyle,M.,
Ferreira,P., Fitzhugh,M., Forrest,C., Funke,R., Gage,D.,
Galagan,J., Gardyna,S., Grant,G., Hagius,B., Heathford,A., Horton,L.,
Howland,J.C., Johnson,R., Jones,C., Krum,L., Karatas,A., Klein,J.,
Lehoczky,J., Lien,C., Locke,K., Macdonald,P., Marquis,N.,
McEwan,P., McGurk,A., McKernan,K., McLaughlin,J., Meldrum,J.,
Morrow,J., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
Peterson,K., Pollara,V., Riley,R., Roy,A., Santos,R., Severy,P.,
Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
Testaye,S., Tirrell,A., Vassiliev,H., Vo,A., Wheeler,J., Wu,X.,
Wyman,D., Ye,W.J., Zimmer,A. and Zody,M.
Direct Submission
Submitted (08-DEC-1999) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Apr 1, 2000 this sequence version replaced g1:6634854.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L3714
Center clone name: 19_B_22
----- Summary Statistics
Sequencing vector: M13; M7815; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 146943 bases at least Q40
Consensus quality: 150092 bases at least Q30
Consensus quality: 151314 bases at least Q20
Insert size: 153000; agarose-fp
Insert size: 152138; sum-of-ctrls
Quality coverage: 4.8 in Q20 bases; agarose-fp
Quality coverage: 4.8 in Q20 bases; sum-of-ctrls
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 10 contigs. The true order of the pieces
* is not known and their order in this sequence record is

```


* consists of 3 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

1 54132: contig of 54132 bp in length
 * 54133 54232: gap of 100 bp
 * 54233 153050: contig of 98818 bp in length
 * 153051 153150: gap of 100 bp
 * 153151 178748: contig of 25598 bp in length.
 Location/Qualifiers
 1. 178748
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="1"
 /clone="RP11-433N2"
 /clone_lib="RPC1-11.2"
 1. 54132
 /note="assembly_fragment:00344
 fragment_chain:1"
 54233. 153050
 /note="assembly_fragment:01015
 fragment_chain:1"
 153151. 178748
 /note="assembly_fragment:01917
 clone_end:77
 vector_side:right"
 BASE COUNT 55375 a 33612 c 33289 g 56272 t 200 others
 ORIGIN

FEATURES

source

misc_feature
 /note="assembly_fragment:00344
 fragment_chain:1"
 54233. 153050
 /note="assembly_fragment:01015
 fragment_chain:1"
 153151. 178748
 /note="assembly_fragment:01917
 clone_end:77
 vector_side:right"
 misc_feature
 /note="assembly_fragment:01015
 fragment_chain:1"
 153151. 178748
 /note="assembly_fragment:01917
 clone_end:77
 vector_side:right"

Query Match 25.4%; Score 178.2; DB 81; Length 178748;
 Best Local Similarity 61.7%; Pred. No. 8.2e-33;
 Matches 360; Conservative 0; Mismatches 203; Indels 20; Gaps 4;

1 gttgtgcaagtgagaaacatggaagaacttaacacacataaaataacaaac 60
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 118439 GTCTTGGCTAGTTGAGAACCCAGAGGACACTCAACAAAACCGTAAATTAACAGAGC 118498
 QY 61 agctctcttgaccattcttagaagaagaagttcaagatccctcttgaagggcactagaa 120
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 118499 AGTTTATCTTACCTGACATC-AGAGAGAGAGAGGAGCAGCATCTCAGAGGCCAAGGAA 118557
 QY 121 gaagaataatctcttggaagaagacacattcaacaaatgaatgagagaccagaagaagatg 180
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 118558 GTGGGAGAGCTGTCCAGACACACACACTCAACAGTGAAGGGAGACCAAGAGAGAAAG 118617
 QY 181 agggactatgtgccaataatgttaactggagatccagggtgttaactagggtgttccaa 240
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 118618 AGGAGACTGTGAGCTGAAGCTTTTACTGTGTGTTCAAGGCGATTACCATGAGAGTTTCCA 118677
 QY 241 tggggaactgttaattgttgatttaagcaagcagagacaaagtccatggagagcattctg 300
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 118678 TGGGAGGTTCTACTGCTGTGATTTTGAGCAAGCAGCAAAATTCATGGGCTCATGCTT 118737
 QY 301 agactgaagaatgactcattgtgcatatctgacagaatctgacgtg----- 349
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 118738 TGAAGGAGAGATGCTGTGACACATGCTGTGCTCAATGCGGTGTGGGGGTGCAGT 118797
 QY 350 attcaagccaaatgagcgtatctagttgtccatagaggtgttaccagagagcagttg 409
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 118798 GGGGAGAGCTCGGTAGCTTGTGATCTGTGCTATGTGAGGTGTGTACACAGAGAGGGA 118857
 QY 410 gtaagtaa-----aaatccctgactgacacatgtgagaaatggaaggtggaagattt 464
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 118858 TTGTGCAAGACAAATATCTGGAATTGACCAATATAAGAACTGGGAAAGCAGAACTG 118917
 QY 465 taaaggtgtcagttgtgactaagaccgtctctctgtatggaataatcaattatttt 524
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 118918 AAAATATGTCAAGATGACCAAGCTCAGTCTGTGATATCAGAAATTAACCTA---TTA 118974

QY 525 aaatgcatagccagacacataaataataagaattaccaca 567
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 118975 AAATGATGCGGAGGACACTTGTGAGAAATCAGACATA 119017

RESULT 15

AC073874 168041 bp DNA HTG 20-AUG-2000
 LOCUS Homo sapiens chromosome 2 clone RP11-586M15, WORKING DRAFT
 DEFINITION AC073874
 ACCESSION AC073874.3 GI:9795984
 VERSION AC073874.3
 KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
 SOURCE human.
 ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 AUTHORS 1 (bases 1 to 168041)
 TITLE The sequence of Homo sapiens clone
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 168041)
 AUTHORS Waterston, R.H.
 TITLE Direct Submission
 JOURNAL Submitted (30-JUN-2000) Genome Sequencing Center, Washington
 University School of Medicine, 4444 Forest Park Parkway, St. Louis,
 MO 63108, USA
 COMMENT On Aug 11, 2000 this sequence version replaced gi:9558670.

COMMENT

----- Genome Center -----
 Center: Washington University Genome Sequencing Center
 Center code: WUGSC
 Web site: http://genome.wustl.edu/gsc/index.shtml
 ----- Project Information -----
 Center project name: H.NH0586M15
 ----- Summary Statistics -----
 Sequencing vector: M13; 100%
 Chemistry: Dye-primer ET; 100% of reads
 Sequencing vector: plasmid; 0%
 Chemistry: Dye-terminator Big Dye; 0% of reads
 Assembly program: Phrap; version 0.990319
 Consensus quality: 159025 bases at least Q40
 Consensus quality: 161851 bases at least Q30
 Consensus quality: 163189 bases at least Q20
 Insert size: 177000; agarose-fp
 Insert size: 166841; sum-of-contigs
 Quality coverage: 4.10 in Q20 bases; agarose-fp
 Quality coverage: 4.42 in Q20 bases; sum-of-contigs
 ----- NOTE: This is a 'working draft' sequence. It currently
 * consists of 13 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.
 1 1595: contig of 1595 bp in length
 * 1596 1695: gap of unknown length
 * 1696 3110: contig of 1415 bp in length
 * 3111 3210: gap of unknown length
 * 3211 6070: contig of 2860 bp in length
 * 6071 6170: gap of unknown length
 * 6171 8472: contig of 2302 bp in length
 * 8473 8572: gap of unknown length
 * 8573 12860: contig of 4288 bp in length
 * 12861 12960: gap of unknown length
 * 12961 18778: contig of 5818 bp in length
 * 18779 18878: gap of unknown length
 * 18879 25944: contig of 7066 bp in length
 * 25945 26044: gap of unknown length
 * 26045 34276: contig of 8232 bp in length
 * 34277 34376: gap of unknown length

